

Обучение систем с дискретным управлением

Научный руководитель – Алисейчик Павел Александрович

Голиков Кирилл Анатольевич

Аспирант

Московский государственный университет имени М.В.Ломоносова,
Механико-математический факультет, Кафедра математической теории
интеллектуальных систем, Москва, Россия

E-mail: golikov.ka@mail.ru

Обучаем систему управлять собой методом проб и ошибок и обобщением их. Весь опыт системы сохраняется в Базе Данных. Система работает в дискретном времени, её статичное состояние описывается координатами нескольких точек на плоскости. Каждый такт на управляющие входы системы можно подавать значения $\{0,1\}$, т.е. включены или выключены приводы, изменяющие состояние системы. Усилия приводов — неизвестные внутренние функции, напрямую управлять ими нельзя. Можно варьировать время запуска и длительность работы каждого привода. Предполагаем, что в приводах системы есть определённые погрешности, состояния, в которые переводят приводы систему, немного отличаются от повторения к повторению. Самым дорогим ресурсом при обучении системы — является время работы системы.

Обучение алгоритма отрабатывается в виртуальной среде для задачи позиционирования роботов разного вида: манипулятор с 2-3 вращательными и призматическими сочленениями и мобильный робот с 2 гусеницами.

Сложности решения задачи в описанных выше ограничениях:

1. За минимальное время произвести системой ценные действия, отражающие возможности и принципы функционирования неизученной системы. Получить управления, хорошо подходящие для интерполяции и предсказаний поведения системы. В рамках задачи позиционирования: определить равномерную плотную решётку целей, получить множество записей в БД движений робота, покрывающих конечными положениями целевые точки из начального положения.

2. Получить функцию интерполяцией по сохранённым записям в БД, которая достаточно точно достигает любые цели на плоскости, а также выбрать новые цели, достичь которые ценно для уточнения интерполяции, чтобы минимизировать возможность промахов.

3. В условиях отсутствия времени на переобучение при изменяющихся функциях усилий приводов системы построить функцию адаптации к этим изменениям. Законы движения робота и изменения в приводах задаются разностными дифференциальными уравнениями, проводя эксперименты с переобучением управления системой важно выяснить границы применимости адаптации для разных видов уравнений.

Написан двухэтапный алгоритм построения равномерной решётки конечных положений, обеспечивает сравнимость и возможность усреднения управлений. Показано, что случайно сгенерированными управлениями, без построения регулярной решётки, за адекватное время обучения не получается хорошей интерполяции.

По решётке приближение к целевым точкам осуществляется методом стохастического градиентного спуска, но показано, что не для всяких двух управлений возможно нахождение промежуточного. Выявляются обособленные кластеры управлений, где спуск работает хорошо. Интерполяция управления строится локальная для ближних конечных положений обеспечивается высокая точность попадания.

Память обширная, предоставляет возможности по построению адекватных предсказаний, но её данные со временем устаревают и перестают верно отображать функционирование системы. Повторить пробы для переобучения времени нет, адаптировать данные нужно "на лету".

После успешной фазы обучения происходит эксплуатация обученной системы. База Данных постоянно дополняется. Так как записи действий в БД сохраняются навсегда, важно не запутаться в одинаковых действиях со смещёнными конечными положениями. Для этого выделяется первое эталонное действие с определённым управлением, при повторениях в указанную запись добавляются фактические смещения базовой точки системы. Чтобы полностью нивелировать смещение, используются статистические методы, для этого все смещения фиксируются и кластеризуются.

Источники и литература

- 1) Яблонский С.В. Введение в дискретную математику. — М.:Высшая школа, 2006.
- 2) Гасанов Э.Э., Кудрявцев В.Б. Интеллектуальные системы. Теория хранения и поиска информации — 2-е изд., испр. и доп. — М.:Юрайт, 2017.
- 3) Саттон Р.С., Барто Э.Г. Обучение с подкреплением — 2-е изд. — М.:БИНОМ.Лаборатория знаний, 2014.